Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Stochastic Processes and their Applications 116 (2006) 1195–1214
stochastic processes and their applications
www.elsevier.com/locate/spa

The proportional hazards regression model with staggered entries: A strong martingale approach

Murray D. Burke[a,*], Dandong Feng[b]

[a] *Department of Mathematics and Statistics, University of Calgary, Calgary, AB, Canada, T2N 1N4*
[b] *St. Mary's University College, Calgary, AB, Canada, T2X 1Z4*

Received 26 August 2004; received in revised form 22 December 2005; accepted 23 February 2006
Available online 20 March 2006

Abstract

The proportional hazards regression model, when subjects enter the study in a staggered fashion, is studied. A strong martingale approach is used to model the two-time parameter counting processes. It is shown that well-known univariate results such as weak convergence and martingale inequalities can be extended to this two-dimensional model. Strong martingale theory is also used to prove weight convergence of a general weighted goodness-of-fit process and its weighted bootstrap counterpart.

MSC: 60G48; 62N01; 62N99

Keywords: Staggered entries; Strong martingales; Proportional hazards regression models

1. Introduction

The Cox proportional hazards regression model has been one of the most studied and used models in statistics. The martingale approach, beginning with the dissertation of [1] and developed by many authors [2], has been very successful in providing a theoretical framework for counting processes, in general, and the Cox model, in particular.

We will study the Cox model with staggered entries. Here one wishes to model hazard as a function of the duration on study. However, one also wishes to sequentially analyze the

* Corresponding address: Department of Mathematics and Statistics, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada, T2N 1N4. Tel.: +1 403 2203959; fax: +1 403 2825150.
E-mail address: burke@math.ucalgary.ca (M.D. Burke).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2006.02.008

clinical trial in calendar time. Thus, a natural two-time parameter stochastic process results. As pointed out in [4] "the standard [one-dimensional] martingale central limit theorem of Rebolledo cannot be applied, at least directly". See also [20,22] and [13]. Bilias et al. [4] proved the weak convergence of the underlying processes using modern empirical process theory. In this paper, a strong martingale approach will be taken to model the processes. For an introduction to strong martingale theory, we refer to [8] and [16]. For two or more time parameters in survival analysis, other martingale approaches have been tried, for example, weak martingales [19]. However, as indicated by Andersen et al. [2], there is no weak martingale central limit theory. For strong martingales, there is the functional central limit theorem of Ivanoff [14] and we will use it to prove weak convergence of our two-time parameter processes.

Although, in practice, one cannot "see" the future in calendar time and thus one cannot "see" the strong past, we use it as a device to prove our limit theorems. In practice, in a sequential approach, an experimenter would use the information up to the current calendar time t to make a decision on whether to stop the experiment and make a statistical conclusion. Since a strong martingale is also 1-martingale, such a sequential approach can be taken.

We will assume that there are potentially infinitely many individuals with entry times denoted by τ_i, failure times T_i, censoring times C_i and $p \times 1$-dimensional vector processes Z_i. Suppose that (τ_i, T_i, C_i, Z_i') are independent and the conditional hazard rate of T_i, given τ_i, C_i, and the covariate process $\{Z_i(u),\ u \le s\}$, is $\lambda_0(s)\ \exp(\beta' Z_i(s))$, the Cox proportional hazards model. For $\widetilde{T}_i = \min\{T_i, C_i\}$, let

$$N_i(t,s) = I[\tau_i + \widetilde{T}_i \le t, \widetilde{T}_i \le s, T_i \le C_i]; \quad Y_i(t,s) = I[\tau_i + s \le t, s \le \widetilde{T}_i];$$
$$A_i(t,s) = \int_0^s Y_i(t,u)\,\lambda_0(u)\exp(\beta' Z_i(u))\,\mathrm{d}u.$$

The two-parameter processes

$$M_i(t,s) = N_i(t,s) - A_i(t,s) \tag{1}$$

and $\sum_{i=1}^n M_i(t,s)$, which will be shown to converge weakly, can be considered as elements of the space $D(S)$, the set of all continuous from above, *lamp* functions with domain S, [23]. A real-valued function x on K is a *lamp* function (has limits along monotone paths), if for each sequence $(t_n, s_n) \in S$, which is monotone in either of the four directions (NE, NW, SE, SW), $\lim_{n\to\infty} x(t_n, s_n)$ exists.

Let t^* be a time such that

$$\lim_{n\to\infty}\inf n^{-1} \sum_{i=1}^n Y_i(t^*, t^*) > 0. \tag{2}$$

Bilias et al. [4] assume (2) and point out that it requires that there be a positive proportion of individuals whose entry times are 0. In this case, we can take $S = \{(t,s) : 0 \le s \le t \le t^*\}$, a triangle. Alternately, one can view a process x on $D(S)$ as a process x^* on $D[0, t^*]^2$ as follows: take $x^*(t,s) = x(t,s)$, if $(t,s) \in S$, and $= x(t, s_1)$, otherwise, where $s_1 = \inf\{s : 0 \le s$ and $(t,s) \in S\}$. If we re-define our processes in this way, then weak convergence will be on the space $D[0, t^*]^2$.

Assumption (2) may not be satisfied. For example, individuals may enter the study according to a continuous distribution with support $(0, t^*)$. However, one can find a point (t_1, s^*), $s^* \le t_1$, such that $\lim\inf_{n\to\infty} n^{-1} \sum_{i=1}^n Y_i(t_1,\ s^*) > 0$. In this latter case, we can take $S = \{(t,s) : 0 \le$

$s \le s^*, t_1 - s^* \le t \le t^* \wedge (s^* - t_1 + s)^+\}$, where $(a \wedge b)^+ = \max\{0, \min\{a, b\}\}$. In general, a set S can be chosen to be convex and to satisfy $\liminf_{n\to\infty} \sum_{i=1}^n n^{-1} Y_i(t, s) > 0$, for all $(t, s) \in S$.

In the next section, we state some definitions and properties of strong martingales. In Section 3 we obtain martingale results for the specific process M_i and for processes based on it. Some of the results would not necessarily be true for all strong martingales but are proven for the present set up. For example, $M_i^2 - A_i$ would not always be a strong martingale (or even a weak martingale). However, the establishment of Lemmas 6 and 7 enables us to prove many subsequent results which have analogues in one-dimensional martingale theory (e.g. Theorems 3–5, and Lenglart-type inequalities). Section 4 establishes the weak convergence of our processes and some likelihood results are proven with martingale methods. In Section 5, partial sum processes, which can be used for model checking, are considered. A $(p+2)$-dimensional strong martingale approach is used to obtain their large sample behavior.

2. Strong martingales

For random elements ξ in $D[0, t^*]^2$, let $\mathcal{F}_x^o = \sigma\{\xi(v) : 0 \le v \le x\}$, for $x = (x_1, x_2)$, where "$\le$" denotes the usual ordering on $[0, t^*]^2$. Let $\mathcal{F}_x = \mathcal{F}_{x_1,x_2}$ denote the right continuous filtration generated by ξ, that is, $\mathcal{F}_x = \bigcap_{v: v_j > x_j, j=1,2} \mathcal{F}_v^o$. Define

$$\mathcal{F}_{x_1}^{(1)} = \bigvee_{0 \le x_2} \mathcal{F}_{x_1,x_2}; \quad \mathcal{F}_{x_2}^{(2)} = \bigvee_{0 \le x_1} \mathcal{F}_{x_1,x_2}, \tag{3}$$

where $\bigvee_{0\le v} \mathcal{F}_{x_1,v}$ is the smallest σ-field generated by the union $\bigcup_{0\le v} \mathcal{F}_{x_1,v}$, and finally, let $\mathcal{F}_x^* = \mathcal{F}_{x_1,x_2}^* = \mathcal{F}_{x_1}^{(1)} \vee \mathcal{F}_{x_2}^{(2)}$.

For any process or function η on $[0, \infty) \times [0, \infty)$ and rectangle

$$D = (t, t'] \times (s, s'], \quad t < t', s < s', \tag{4}$$

the value of η on D is defined as

$$\eta(D) = \eta(t', s') - \eta(t', s) - \eta(t, s') + \eta(t, s). \tag{5}$$

Let ξ be a random element of $D[0, t^*]^2$ and $\mathcal{F}_x^o$ be any filtration to which ξ is adapted. Then, ξ is a *strong martingale* if it is integrable, $E|\xi(t, s)| < \infty$ for all $(t, s) \in [0, t^*]^2$, and $E[\xi(D)|\mathcal{F}_{t,s}^*] = 0$, where D is defined in (4). Here $\mathcal{F}_{t,s}^*$ denotes the information contained in the L-shaped region $L_x = \{(u, v) \in [0, t^*]^2 : u \le t, \text{or } v \le s\}$. We quote the following version of Ivanoff's functional central limit theorem for strong martingales in the space $D_d = D[0, t^*]^d$.

Theorem 1 ([14]). *For $n = 1, 2 \ldots; \infty$, let $\xi_n(\cdot) \in D_d = D[0, t^*]^d$ be a strong martingale. If for each t, $\{\xi_n(t)\}_n$ is uniformly integrable, the finite dimensional distributions of ξ_n converge to the corresponding ones of ξ_∞, and ξ_∞ has a version with almost-sure continuous sample paths, then*

$$\xi_n \to_D \xi_\infty.$$

The point of this result is that for strong martingales, tightness automatically follows when trying to establish weak convergence to a process with almost-sure continuous sample paths.

3. The strong martingale M_i

Let $\mathcal{F}_{t,s} = \bigvee_{i=1}^{n} \mathcal{F}_{t,s}^{i}$ be the smallest σ-field containing

$$\begin{aligned}\mathcal{F}_{t,s}^{i} = \sigma\{&I_{[\tau_i \le v]}, \tau_i I_{[\tau_i \le v]}, I_{[T_i \le u \wedge C_i \wedge (v-\tau_i)^+]}, T_i I_{[T_i \le u \wedge C_i \wedge (v-\tau_i)^+]},\\ &I_{[C_i \le u \wedge T_i \wedge (v-\tau_i)^+]}, C_i I_{[C_i \le u \wedge T_i \wedge (v-\tau_i)^+]},\\ &Z_i(u \wedge (v-\tau_i)^+) I_{[\tau_i \le v]},\ 0 \le u \le s,\ 0 \le v \le t\},\end{aligned}$$

$1 \le i \le n$, where each Z_i is piecewise constant with probability one and $v^+ = \max\{v, 0\}$. Then, the filtration $\{\mathcal{F}_{t,s} : t, s \ge 0\}$ is right-continuous. See [10] for a proof. Let $D = (t_0, t_0'] \times (s_0, s_0']$, where $0 \le t_0 < t_0'$ and $0 \le s_0 < s_0'$, be an arbitrary rectangle. We have

Theorem 2. *The process $M_i(t, s)$, defined by (1) is a strong martingale relative to $\{\mathcal{F}_{t,s}^{*}\}$.*

The proofs of the results of this section are in Section 6. In order to reflect measurability considerations of our integrands and that we observe the value $Z_i(u)$ when $Y_i(t, u) = 1$, that is, when the individual is observed to be at risk at calendar time t, we will write

$$\int_{(0,s]} Z_i(u)\, M_i(t, \mathrm{d}u) = \int_{(0,s]} Y_i(t, u) Z_i(u) M_i(t, \mathrm{d}u).$$

Since our integrals are with respect to one variable only, we introduce the notion of 2-predictability as follows. Let $\mathcal{P}_t$ denote the σ-field of $\mathcal{F}_{t,s}$ generated by subsets of the form $\{0\} \times A$, $A \in \mathcal{F}_{t,0}$; $(a, b] \times A$, $0 \le a < b < \infty$, $A \in \mathcal{F}_{t,a}$. We call $\mathcal{P}_t$ the 2-predictable σ-field. A process $H(t, s)$ is said to be 2-*predictable*, with respect to $\{\mathcal{F}_{t,s} : t, s \ge 0\}$ if, for each t, as a map from $[0, \infty) \times \Omega$ to R, it is measurable with respect to $\mathcal{P}_t$.

A 2-predictable process can be generated by simple 2-predictable processes of the form $c_0 I_{[0] \times A_0} + \sum_{j=1}^{k} c_j I_{(a_j, b_j] \times A_j}$, where $A_0 \in \mathcal{F}_{t,0}$, $A_j \in \mathcal{F}_{t,s}$, and c_j for $j = 1, 2, \ldots, k$, are constants. As in the one-dimensional case, the integral $L(t, s) = \int_{(0,s]} H(t, u)\, M_i(t, \mathrm{d}u)$ can be well-defined as an $\mathcal{F}_{t,s}$-adapted process. If $H(t, s)$ is measurable with respect to $\mathcal{F}_{t,s}$ and left continuous in s for each t, then H is 2-predictable.

Define

$$\begin{aligned} J_{ij}(t, s) &= \int_0^s H_{ij}(u) Y_i(t, u) M_i(t, \mathrm{d}u), \quad j = 1, 2,\\ A_i^{*}(t, s) &= \int_0^s H_{i1}(u) H_{i2}(u) A_i(t, \mathrm{d}u),\\ M_i^{*}(t, s) &= J_{i1}(t, s) J_{i2}(t, s) - A_i^{*}(t, s). \end{aligned} \tag{6}$$

Theorem 3. *If the process $Y_i(t, u) H_{i1}(u)$ is 2-predictable, where H_{i1} is a univariate process, then $J_{i1}(t, s)$ is a strong martingale with respect to $\mathcal{F}_{(t,s)}^{*}$.*

Theorem 4. *If the processes $Y_i(t, u) H_{ij}(u)$ are 2-predictable, $j = 1, 2$, where H_{ij} are univariate processes, then $M_i^{*}(t, s)$ is a strong martingale with respect to $\mathcal{F}_{(t,s)}^{*}$.*

Theorem 5. *Suppose that $(H_{i1}, H_{i2}, Y_i, M_i)$, $i = 1, 2, \ldots, n$, are independent and the conditions of Theorem 4 hold. Then, the processes $\sum_{i=1}^{n} J_{ij}$, $j = 1, 2$, and $M^{*}(t, s) = \sum_{i=1}^{n} M_i^{*}(t, s)$ are strong martingales with respect to $\mathcal{F}_{(t,s)}^{*}$.*

If we take $H_{ij} \equiv 1$, $j = 1, 2$, $1 \le i \le n$, in Theorem 5, we have $M_i^2 - A_i$ is a strong martingale relative to $\mathcal{F}^*_{t,s}$. Theorems 4 and 5 are not true for a general strong martingale M_i [12]. However, for M_i defined by (1), they are. This is because of the following two lemmas.

Lemma 6. *Assume the conditions of Theorem 4. If $t \le t'$, $s \le s'$, then $[J_{i1}(t, s') - J_{i1}(t, s)][J_{i2}(t', s) - J_{i2}(t, s)] = 0$, for $i = 1, 2, \ldots, n$.*

Lemma 7. *Under the conditions of Theorem 4, for the increment of M_i^* over rectangle $D = ((t, s), (t', s')]$, we have $E[M_i^*(D)|\mathcal{F}^*_{t,s}] = E[J_{1i}(D)J_{2i}(D) - A_i^*(D)|\mathcal{F}^*_{t,s}]$, a.s.*

The first lemma of the next two is adapted without proof from Lemma 3, statement (12), of [17].

Lemma 8. *Let $M(t, s)$ be a square integrable strong martingale. If its compensator $A(t, s)$ is continuous with probability one, then for any $(t^*, s^*) \in [0, \infty) \times [0, \infty)$, $\varepsilon, \eta > 0$,*

$$P\left[\sup_{(t,s)\le(t^*,s^*)} |M(t, s)| \ge \varepsilon\right] \le P[A(t^*, s^*) > \eta] + \varepsilon^{-2} E[A(t^*, s^*) \wedge 2\eta].$$

Lemma 9. *Let $N(t, s) = \sum_{i=1}^n N_i(t, s)$, $A(t, s) = \sum_{i=1}^n A_i(t, s)$ and $M = N - A$. If the process $Y_i(t, u)H_{i1}(u)$ is 2-predictable, then for any $\varepsilon, \eta \ge 0$,*

$$P\left[\sup_{(t,s)\le(t^*,t^*)} |N(t, s)| \ge \varepsilon\right] \le \frac{\eta}{\varepsilon^2} + 2P\left[A(t^*, t^*) \ge \left(\frac{\eta}{4} \wedge \frac{\varepsilon}{2}\right)\right]; \quad \textit{and}$$

$$P\left[\sup_{(t,s)\le(t^*,t^*)} \left(\sum_{i=1}^n J_{i1}(t, s)\right)^2 \ge \varepsilon\right] \le \frac{\eta}{\varepsilon^2} + 2P\left[\sum_{i=1}^n \int_0^{t^*} H_{i1}^2(u) A_i(t^*, \mathrm{d}u) \ge \frac{\eta}{4}\right].$$

4. Estimation and asymptotic theory

First we consider the likelihood function

$$L(\beta; t, s) = \prod_{i=1}^n \prod_{0\le u\le s} \left[\frac{\exp(\beta' Z_i(u))}{\sum_{l=1}^n Y_l(t, u)\exp(\beta_l' Z(u))}\right]^{\triangle N_i(t,u)} \tag{7}$$

and the score process

$$\begin{aligned} U(\beta; t.s) &= \sum_{i=1}^n \int_0^s [Y_i(t, u)Z_i(u) - \overline{Z}(\beta; t, u)]N_i(t, \mathrm{d}u) \\ &= U_2(\beta; t, s) - \int_0^s \overline{Z}(\beta; t, u)U_1(\beta; t, \mathrm{d}u), \end{aligned} \tag{8}$$

where $U_1(\beta, t, s) = \sum_{i=1}^{n} M_i(\beta; t, s)$, $U_2(\beta; t, s) = \sum_i \int_0^s Y_i(t, u) Z_i(u) M_i(\beta; t, \mathrm{d}u)$ and

$$\overline{Z}(\beta; t, u) = \frac{\sum_{l=1}^{n} Y_l(t, u) Z_l(u) \exp(\beta_l' Z(u))}{\sum_{l=1}^{n} Y_l(t, u) \exp(\beta_l' Z(u))}.$$

As pointed out by [4], $U(\beta; t, t)$ is the partial likelihood score process $\nabla_\beta \log L(\beta; t, t)$ of the Cox partial likelihood. Hwang [13] showed that for fixed $t_1, \ldots, t_k$, $U_j(\beta, t_1, s)$, $U_j(\beta, t_2, s) - U_j(\beta, t_1, s), \ldots, U_j(\beta, t_k, s) - U_j(\beta, t_{k-1}, s)$ are orthogonal one-dimensional martingales (in s), where U_j is the jth component of U. He studied the asymptotic behavior of $U_1(\beta; t, t)$ and showed that properly normalized $\hat{\beta}_1 - \beta_{10}$ behaves like Brownian motion in an information-based clock. Our Theorems 2 and 3, above, establish that the processes U_1 and U_2 are strong martingales, which is a stronger result.

Analogous to the notation in [3], we define:

$$S^{(k)}(\beta, t, s) = n^{-1} \sum_{i=1}^{n} Z_i(s)^{\otimes k} Y_i(t, s) \exp(\beta' Z_i(s)), \quad k = 0, 1, \text{ and } 2.$$

$$E(\beta, t, s) = \frac{S^{(1)}(\beta, t, s)}{S^{(0)}(\beta, t, s)}; \quad V_n(\beta, t, s) = \frac{S^{(2)}(\beta, t, s)}{S^{(0)}(\beta, t, s)} - E(\beta, t, s)^{\otimes 2},$$

where, for column p-vectors a and b, $a \otimes b = ab'$, a matrix, and $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$, $a^{\otimes 2} = aa'$. Also, let $\|B\|$ denote the maximum of the components of the matrix (or vector) B.

We will assume the following conditions:

Condition C1. There exists a t^* such that λ_0 is bounded on $[0, t^*]$.

Condition C2. There exists a constant c such that the total variation $|Z_i(0)| + \int_0^{t^*} |\mathrm{d}Z_i(u)| \le c$, where the first $|\cdot|$ denotes the L_1-norm for a p-dimensional vector and the second one the L_1-type total variation for a p-dimensional vector function.

Condition C3. Let $\mathcal{B}$ be a neighborhood of β_0. There exist $s^{(0)}$, a p-vector $s^{(1)}$ and a $p \times p$-matrix $s^{(2)}$ defined on $\mathcal{B} \times S_*$, where $S_* = [(t, s) : t_* \le s \le t \le t^*]$, such that, $\lim_{n\to\infty} E[S^{(k)}(\beta, t, s)] = s^{(k)}(\beta, t, s)$, and the functions $s^{(0)}$, $s^{(1)}$ and $s^{(2)}$ are of bounded variations and left-continuous in s for each t.

Condition C4. Let $e = s^{(1)}/s^{(0)}$ and $V = s^{(2)}/s^{(0)} - e^{\otimes 2}$. Then, for all $\beta \in \mathcal{B}$ and $(s, t) \in S_*$,

$$(\partial/\partial\beta) s^{(0)}(\beta, t, s) = s^{(1)}(\beta, t, s); \quad (\partial^2/\partial\beta^2) s^{(0)}(\beta, t, s) = s^{(2)}(\beta, t, s)$$

Condition C5. Let e be defined as above, and

$$K_n(\beta_0, t, s) = \frac{\sum_{i=1}^{n} E[Z_i(s) Y_i(t, s) \exp(\beta_0' Z_i(s))]}{\sum_{i=1}^{n} E[Y_i(t, s) \exp(\beta_0' Z_i(s))]}. \tag{9}$$

Then $\sup_{0\le t\le t^*} \int_0^t [K_n(\beta_0, t, s) - e(\beta_0, t, s)]^2 \mathrm{d}s \to 0$.

Condition C6. The function $s^{(0)}$ is bounded away from 0 on S_*; for $k = 0, 1, 2$, the family of $s^{(k)}(\cdot, t, s)$, $(t, s) \in S_*$ is equicontinuous at β_0.

Condition C7. The matrix $\Sigma(\beta_0, t, s) = \int_0^s V(\beta_0, t, u) s^{(0)}(\beta_0, t, u) \lambda_0(u) \mathrm{d}u$ is positive definite.

Theorem 10. *Suppose conditions C1, C2, C3 and C6 are satisfied. Then, the vector* $n^{-1/2}(U_1(\beta_0, t, s), U_2(\beta_0, t, s))$ *converges in distribution, in the space* $D^{p+1}(S_*)$*, to the vector* $(\xi_1(t, s), \xi_2(t, s))$*, a vector-valued Gaussian process with zero mean and covariance function*

$$\begin{aligned} E[\xi_1(t_1, s_1)\xi_1(t_2, s_2)] &= \int_0^{s_1 \wedge s_2} s^{(0)}(\beta_0, t_1 \wedge t_2, u)\lambda_0(u)\mathrm{d}u, \\ E[\xi_2(t_1, s_1)\xi_2'(t_2, s_2)] &= \int_0^{s_1 \wedge s_2} s^{(2)}(\beta_0, t_1 \wedge t_2, u)\lambda_0(u)\mathrm{d}u, \\ E[\xi_1(t_1, s_1)\xi_2(t_2, s_2)] &= \int_0^{s_1 \wedge s_2} s^{(1)}(\beta_0, t_1 \wedge t_2, u)\lambda_0(u)\mathrm{d}u. \end{aligned} \tag{10}$$

Proof of theorem 10. Both U_1 and U_2 are strong martingales with expected values zero. By the multivariate central limit theorem and regularity condition *C3*, the finite-dimensional distributions of $n^{-\frac{1}{2}}U_1(\beta_0, t, s)$ and $n^{-\frac{1}{2}}U_2(\beta_0, t, s)$ converge to a multivariate normal with mean zero and covariance matrix specified in (10). To complete the proof of weak convergence, we must prove tightness of the vector-valued processes.

To show tightness of $n^{-\frac{1}{2}}(U_1, U_2)$, it suffices to prove it component-wise. Using Theorem 1, we need only show that there are versions of Gaussian processes with covariance (10) having continuous sample paths, almost surely. The function $\int_0^{s_1 \wedge s_2} s^{(0)}(\beta_0, t_1 \wedge t_2, u)\lambda_0(u)\mathrm{d}u$ and the diagonal components of $\int_0^{s_1 \wedge s_2} s^{(2)}(\beta_0, t_1 \wedge t_2, u)\lambda_0(u)\mathrm{d}u$ are nondecreasing functions in (t, s), so that

$$F_0(t, s) = \frac{\int_0^s s^{(0)}(\beta_0, t, u)\lambda_0(u)\mathrm{d}u}{\int_0^{t^*} s^{(0)}(\beta_0, t^*, u)\lambda_0(u)\mathrm{d}u}; \quad F_j(t, s) = \frac{(\int_0^s s^{(2)}(\beta_0, t, u)\lambda_0(u)\mathrm{d}u)_{jj}}{(\int_0^{t^*} s^{(2)}(\beta_0, t^*, u)\lambda_0(u)\mathrm{d}u)_{jj}},$$

$j = 1, 2, \ldots, p$, are continuous distribution functions on $[0, t^*] \times [0, t^*]$. Hence for $j = 0, 1, \ldots, p$, there exists independent random vectors, with distribution function F_j, whose empirical process converges weakly to a Gaussian process (a so-called Kiefer process) with continuous sample paths and with zero mean and covariance $F_j(t_1 \wedge t_2, s_1 \wedge s_2) - F_j(t_1, s_1)F_j(t_2, s_2)$, [9]. Then, ξ_1 and the components ξ_{2j} of ξ_2 represent the "Wiener" part of the Kiefer processes and hence it has a version with continuous sample paths with probability one. By Theorem 1, the components of $n^{-\frac{1}{2}}(U_1, U_2)$ converge in distribution to those of $n^{-\frac{1}{2}}(\xi_1, \xi_2)$ and hence each is tight. The covariance (10) can be deduced directly from Theorems 3–5. □

Lemma 11. *For* K_n *is defined by* (9)*, let*

$$\widetilde{U}(\beta_0; t, s) = U_2(\beta_0; t, s) - \sum_{i=1}^{n} K_n(\beta_0; t, u)M_i(t, u). \tag{11}$$

Then, under Conditions C1–C7, the score process $n^{-1/2}U(\beta_0; t, s)$ *of* (8) *and* $n^{-1/2}\widetilde{U}(\beta_0; t, s)$ *have the same asymptotic distribution.*

See (2.8) of [4] for a proof. We obtain:

Theorem 12. *Under Conditions C1–C7* $n^{-1/2}U \to_D \xi$*, in the space* $D^p(S_*)$*, where* $U = U(\beta_0; t, s)$ *is defined by* (8) *and* ξ *is a vector-valued Gaussian process and mean 0 and covariance*

$$E[\xi(t_1, s_1)\, \xi'(t_2, s_2)] = \Sigma(\beta_0, t_1 \wedge t_2, s_1 \wedge s_2). \tag{12}$$

Proof of theorem 12. By Lemma 11, we need to show the weak convergence of $n^{-1/2}\widetilde{U}(\beta_0, t, s)$ to ξ. (It is sufficient to show the weak convergence of $n^{-1/2}\widetilde{U}(\beta_0, t, s)$ to ξ componentwise. Without loss of generality, we assume that the Z_i are scalars.) By Skorohod–Dudley–Wichura strong representation theorem [21], there exists versions $n^{-1/2}\sum_{i=1}^{n}\widetilde{M}_i$ and $\widetilde{\xi}_1$ whose distributions are the same as those of $n^{-1/2}\sum_{i=1}^{n} M_i$ and ξ_1 respectively such that

$$\sup_{(t,s)\in D^*}\left|n^{-1/2}\sum_{i=1}^{n}\widetilde{M}_i(t,s)-\widetilde{\xi}_1(t,s)\right| \to 0 \text{ a.s.} \tag{13}$$

Let (with β_0 suppressed)

$$U_1^K(t,s) = n^{-1/2}\int_0^s K_n(t,u)\sum_{i=1}^{n} M_i(t,\mathrm{d}u); \quad \xi_1^e(t,s) = \int_0^s e(t,u)\xi_1(t,\mathrm{d}u);$$

$$\widetilde{U}_1^K(t,s) = n^{-1/2}\int_0^s K_n(t,u)\sum_{i=1}^{n} \widetilde{M}_i(t,\mathrm{d}u); \quad \widetilde{\xi}_1^e(t,s) = \int_0^s e(t,u)\widetilde{\xi}_1(t,\mathrm{d}u).$$

By regularity condition $C3$, $K_n(t,s)$ is of bounded variation in s uniformly in t, then (13) implies $\sup_{(t,s)\in D^*}|\widetilde{U}_1^K(t,s)-\widetilde{\xi}_1^e(t,s)| \to 0$ a.s., by (A.2) of Lemma A.3. of [4]. Hence

$$U_1^K =_{\mathcal{D}} \widetilde{U}_1^K \to \widetilde{\xi}_1^e(t,s) = \int_0^s e(t,u)\widetilde{\xi}_1(t,\mathrm{d}u) =_{\mathcal{D}} \int_0^s e(t,u)\xi_1(t,\mathrm{d}u).$$

Thus, U_1^K is tight. Since the finite-dimensional distributions converge to those of ξ_1^e we have $U_1^K \to_D \xi_1^e$. By Lemma 6 of [23], since U_2 has a continuous limiting process, $\widetilde{U}(\beta_0; t, s) = U_2(t,s) - U_1^K(t,s)$ is tight. Since the finite-dimensional distributions of $\widetilde{U}(\beta_0; t, s)$ converge to those of ξ, $n^{-1/2}\widetilde{U} = n^{-1/2}(U_2 - U_1^K) \to_D \xi$.

If the entry times τ_i are independent of (T_i, C_i, Z_i), then the function K_n of (9), and hence the functions e and V of condition C4, are independent of t. In this case, $\widetilde{U}(\beta_0; t, s)$ is a strong martingale and Theorem 1 can be directly applied. The form of the covariance (12) follows immediately from the strong martingale properties. For points (t,s) and (t',s'), with $s \le s'$, let D_1 be the rectangle $(0,t'] \times (s,s']$. Since

$$\begin{aligned}&E[\widetilde{U}(\beta_0; t, s)(\widetilde{U}'(\beta_0; t', s') - \widetilde{U}'(\beta_0; t', s))|\mathcal{F}^*_{0,s}]\\ &\quad = E[\widetilde{U}(\beta_0; t, s)\widetilde{U}'(\beta_0; D_1)|\mathcal{F}^*_{0,s}] = \widetilde{U}(\beta_0; t, s)E[\widetilde{U}'(\beta_0; D_1)|\mathcal{F}^*_{0,s}] = 0 \text{ a.s.,}\end{aligned}$$

then $E\widetilde{U}(\beta_0; t, s)\widetilde{U}'(\beta_0; t', s') = E\widetilde{U}(\beta_0; t, s)\widetilde{U}'(\beta_0; t, s)$. Also, on letting $D_2 = (t' \wedge t, t' \vee t] \times (0, s]$,

$$\begin{aligned}&E[\widetilde{U}(\beta_0; t' \wedge t, s)(\widetilde{U}'(\beta_0; t' \vee t, s) - \widetilde{U}'(\beta_0; t' \wedge t, s))|\mathcal{F}^*_{t'\wedge t,0}]\\ &\quad = \widetilde{U}(\beta_0; t' \wedge t, s)E[\widetilde{U}'(\beta_0; D_2)|\mathcal{F}^*_{t'\wedge t,0}] = 0 \text{ a.s.}\end{aligned}$$

Hence $E\widetilde{U}(\beta_0; t, s)\widetilde{U}'(\beta_0; t', s) = E\widetilde{U}(\beta_0; t' \wedge t, s)\widetilde{U}'(\beta_0; t' \wedge t, s)$, which, when divided by n, has limit (12).

If the entry times τ_i depend on (T_i, C_i, Z_i), then the function K_n of (9) would, in general, depend on both t and s, and $\widetilde{U}(\beta_0; t, s)$ would not be a strong martingale. However, using Theorem 10 and the fact that U_1 and U_2 are strong martingales, the limiting covariance (12) can be obtained from the following: for $s \le s'$ and with the argument $\beta = \beta_0$ suppressed,

$$EU_2(t,s)U_1^K(t',s') = \sum_{i=1}^n \int_0^s K_n(t',u)ES^{(1)}(t \wedge t', u)\lambda_0(u)\mathrm{d}u;$$

$$EU_1^K(t,s)U_2(t',s') = \sum_{i=1}^n \int_0^s K_n(t,u)ES^{(1)}(t \wedge t', u)\lambda_0(u)\mathrm{d}u;$$

$$EU_1^K(t,s)U_1^K(t',s') = \sum_{i=1}^n \int_0^s K_n(t,u)K_n(t',u)ES^{(0)}(t \wedge t', u)\lambda_0(u)\mathrm{d}u. \quad \square$$

Lemma 13. *Under the regularity conditions C1–C7, if $\widehat{\beta}_n(t,s)$ is a maximum of the concave function (7), then* $\sup_{(t,s)\in S_*} \|\widehat{\beta}_n(t,s) - \beta_0\| \to_{\text{a.s.}} 0$.

See [4] for a proof.

Theorem 14. *If $\widehat{\beta}$ is any consistent estimator of β_0, then*

$$\sup_{(t,s)\le(t^*,t^*)} \left\| n^{-1}\mathcal{I}(\widehat{\beta},t,s) - \Sigma(\beta_0;t,s) \right\| \to_P 0,$$

as $n \to \infty$, where $\mathcal{I}(\widehat{\beta},t,s) = \sum_{i=1}^n \int_0^s V_n(\widehat{\beta},t,u)N_i(t,\mathrm{d}u)$.

Using Pollard's uniform strong law of large numbers [18], Lemma A.3 of [4], and standard arguments [3], conditions *C1–C4* imply the result. In the case that the entry times τ_i are independent of (T_i, C_i, Z_i), the function $V(\beta,t,s)$ (see condition *C4*) is independent of t. Thus, $\int_0^s V(\beta_0,t,u)n^{-1}U_1(t,\mathrm{d}u)$ is a strong martingale and Lemma 9 can be used as Lenglart's inequality was used in [3]. Standard arguments [4] lead to

Theorem 15. *Under the regularity conditions C1–C7, $\{\sqrt{n}(\widehat{\beta}(t,s) - \beta_0), t_* \le s \le t \le t^*\}$ converges weakly to a Gaussian process η which has mean zero and covariance*

$$E(\eta(t_1,s_1)\eta'(t_2,s_2)) = \Sigma^{-1}(\beta_0,t_1,s_1)\Sigma(\beta_0,t_1 \wedge t_2, s_1 \wedge s_2)\Sigma^{-1}(\beta_0,t_2,s_2).$$

An estimator of the baseline cumulative hazard function, given β, is

$$\widehat{\Lambda}(\beta,t,s) = \int_0^s \left[\sum_i Y_i(t,u)\mathrm{e}^{\beta' Y_i(t,u)Z_i(u)}\right]^{-1} \left(\sum_i N_i(t,\mathrm{d}u)\right). \tag{14}$$

Under regularity conditions *C1–C7*, in the space $D(\widetilde{S}_*)$, $\sqrt{n}\{\widehat{\Lambda}(\widehat{\beta}(t,t),t,s) - \Lambda_0(s)\}$ converges in distribution to a Gaussian random field with mean 0 and covariance function

$$\begin{aligned} &\int_0^{s_1 \wedge s_2} \frac{\lambda_0(u)}{s^{(0)}(t_1 \vee t_2)}\mathrm{d}u + Q'(t_2,s_2)\Sigma^{-1}(t_1 \vee t_2, t_1 \vee t_2)Q(t_1,s_1) \\ &\quad - Q'(t_1,s_1)\Sigma^{-1}(t_1,s_1)\{Q(t_1 \wedge t_2, s_1\delta_1 \vee s_2\delta_2) \\ &\quad - Q(t_1 \vee t_2, s_1\delta_1 \vee s_2\delta_2)\}\Sigma^{-1}(t_2,s_2)Q(t_2,s_2). \end{aligned} \tag{15}$$

where $Q(t,s) = \int_0^s \{s^{(1)}(t,u)/s^{(0)}(t,u)\}\lambda_0(u)\mathrm{d}u$; $\delta_1 = I_{[t_1 \le t_2]}$; $\delta_2 = I_{[t_1 > t_2]}$. Especially, if $t_1 = t_2 = t$, the covariance function (15) simplifies to $E[\varsigma(t,s_1)\varsigma(t,s_2)] = \int_0^{s_1 \wedge s_2} \frac{\lambda_0(u)}{s^{(0)}(t)}\mathrm{d}u +$

$Q'(t, s_2)\Sigma^{-1}(t, t)Q(t, s_1)$. The function Q is independent of t, if the τ_i are independent of (T_i, C_i, Z_i).

Remark. To apply these asymptotic results, the value of the total number of individuals n must be specified in advance. Alternately, if the entry times τ_i, $(1, 2, \ldots, n)$, come from a known distribution function G, then we can replace the factor n in the above results with $\overline{Y}(t, 0) = \sum_{i=1}^{n} Y_i(t, 0)$, which is the empirical distribution function of the τ_i, $(1, 2, \ldots, n)$. We then obtain, under Conditions *C1* to *C7*,

$$\left(\overline{Y}(t, 0)\right)^{-1/2} U(\beta_0, t, s) \to_D (G(t))^{-1/2}\, \xi(t, s)$$
$$\left(\overline{Y}(t, 0)\right)^{1/2} (\widehat{\beta}_n(t, s) - \beta_0) \to_D (G(t))^{1/2}\eta(t, s),$$

since $\sup_t |n^{-1} \sum_{i=1}^{n} Y_i(t, 0) - G(t)| \to 0$, as $n \to \infty$, by the Glivenko–Cantelli theorem.

5. Model checking

We wish to consider processes useful in testing the underlying proportions hazards model. Assume $(Z_i,\ C_i,\ T_i,\ \tau_i)$, $(i = 1, 2, \ldots, n)$ are independent and identically distributed with distribution function F_0 and that given Z_i (assumed to be time independent), the random variables τ_i, T_i and C_i are independent.

Consider the weighted $(p + 2)$ parameter cumulative martingale-residual process:

$$\widehat{\psi}_n^f(z, t, s) = n^{-1/2} \sum_{i=1}^{n} f(Z_i)\, I_{[Z_i \le z]} \widehat{M}_i(t, s), \tag{16}$$

with $z \in [a, b] \equiv \prod_{l=1}^{p} [a_l, b_l]$, where f is a given weight function,

$$\widehat{M}_i(t, s) = N_i(t, s) - \int_0^s Y_i(t, u) \mathrm{e}^{\widehat{\beta}(t,t)' Z_i} \widehat{\Lambda}(t, \mathrm{d}u),$$

and $\widehat{\Lambda}$ is defined by (14). Lin et al. [15] consider this general process in the case $\tau_i \equiv 0$. For M_i is defined by (1), let

$$\psi_n^f(z, t, s) = n^{-1/2} \sum_{i=1}^{n} f(Z_i) I_{[Z_i \le z]} M_i(t, s). \tag{17}$$

We will denote ψ_n^f by ψ_n^1 when $f(z) = 1$, for all z. Let

$$\begin{aligned}\mathcal{F}_{z,t,s}^{i} = \sigma\{ & I_{[\tau_i \le v]}, \tau_i I_{[\tau_i \le v]}, I_{[T_i \le u \wedge C_i \wedge (v - \tau_i)^+]}, T_i I_{[T_i \le u \wedge C_i \wedge (v - \tau_i)^+]}, \\ & I_{[C_i \le u \wedge T_i \wedge (v - \tau_i)^+]}, C_i I_{[C_i \le u \wedge T_i \wedge (v - \tau_i)^+]}, Z_i I_{[Z_i \le z', \tau_i \le v]}, \\ & I_{[Z_i \le z', \tau_i \le v]},\ 0 \le u \le s, 0 \le v \le t, a \le z' \le z\},\end{aligned}$$

and define $\mathcal{F}_{z,t,s} = \bigvee_{i=1}^{n} \mathcal{F}_{z,t,s}^{i}$. Then, ψ_n^f and $\widehat{\psi}_n^f$ are measurable with respect to $\mathcal{F}_{z,t,s}$. Define

$$\mathcal{F}_{z,t,s}^{*} = \mathcal{F}_x^{*} = \bigvee_{j=1}^{p+2} \mathcal{F}_{x_j}^{(j)}, \tag{18}$$

as in Section 2, where $x = (x_1, x_2, \ldots, x_{p+2}) = (z, t, s)$ and (see (3)) $\mathcal{F}_{x_j}^{(j)} = \bigvee_{x_\ell : \ell \ne j} \mathcal{F}_x$, $j = 1, 2, \ldots, p + 2$.

Theorem 16. *The process* ψ_n^f *is a* $(p+2)$*-dimensional strong martingale with respect to* $\mathcal{F}^*_{z,t,s}$ *and*

$$\psi_n^f \to_D \psi_\infty,$$

in the space $D([a,b)^p \times S_*)$*, where* S_* *is defined in Condition (C3),* ψ_∞ *is a Gaussian process with mean zero and covariance function:*

$$E\psi_\infty(z_1,t_1,s_1)\psi_\infty(z_2,t_2,s_2) = \Upsilon_2(z_1 \wedge z_2, t_1 \wedge t_2, s_1 \wedge s_2),$$

where $\Upsilon_j(z,t,s) = E\left[[f(Z_i)]^j I_{[Z_i \le z]} \int_0^s Y_i(t,u) e^{\beta_0' Z_i} \lambda_0(u) du\right]$.

When comparing the asymptotic distribution of $\widehat{\psi}_n^f$ with that of ψ_n^f, additional terms appear. In this case there are two types: one for the estimated parameters, and one because the baseline hazard function is unknown and must be estimated. Similar to the one-time parameter representation case of [15], we have the following representation:

$$\widehat{\psi}_n(z,t,s) = \psi_n^f(z,t,s) - \int_0^t g(\beta_0,z,t,u) n^{-1/2} \sum_{i=1}^n M_i(t,du) \\ - W_n(z,t,s)\widetilde{U}(\beta_0,t,t), \tag{19}$$

where $\widetilde{U}$ is defined by (11), $g(\beta_0,z,t,u) = \Upsilon_1(z,t,s)/\Upsilon_0(b,t,s)$,

$$W_n(z,t,s) = n^{-1}\mathcal{J}^{-1}(\beta_0,t) \sum_{k=1}^n \int_0^s Y_k(t,u) e^{\beta_0' Z_k} f(Z_k) I_{[Z_k \le z]} \\ \times \{Z_k - K(\beta_0,t,u)\} \lambda_0(u) du,$$

$K = K_n$ (independent of n) is defined by (9) and $\mathcal{J}(\beta,t)$ is minus the derivative (with respect to β) matrix of $U(\beta,t,t)$. We have

Theorem 17. $\widehat{\psi}_n \to_D \varsigma$*, in the space* $D([a,b)^p \times \widetilde{S}_*)$*, where* ς *is a mean zero Gaussian process with representation*

$$\varsigma(z,t,s) = \psi_\infty^f(z,t,s) - \int_0^t g(\beta_0,z,t,u)\, \psi_\infty^1(b,t,du) - W(z,t,s)\xi(t,t),$$

where $W(z,t,s) = \lim_{n\to\infty} EW_n(z,t,s)$*, and* $(\psi_\infty^f, \psi_\infty^1, \xi)$ *are jointly Gaussian with mean zero and covariance calculated from*

$$E\psi_\infty^f(z_1,t_1,s_1)\psi_\infty^1(z_2,t_2,s_2) = \Upsilon_1(z_1 \wedge z_2, t_1 \wedge t_2, s_1 \wedge s_2),$$
$$E\psi_\infty^1(z_1,t_1,s_1)\psi_\infty^1(z_2,t_2,s_2) = \Upsilon_0(z_1 \wedge z_2, t_1 \wedge t_2, s_1 \wedge s_2),$$
$$E\psi_\infty^f(z,t_1,s)\xi(t_2,t_2) = E\int_0^s (Z_i - K(\beta_0,t_2,u)) f(Z_i) I_{[Z_i \le z]} Y_i(t_1 \wedge t_2, u) e^{\beta_0' Z_i} \Lambda_0(du)$$
$$E[\xi(t_1,t_2)]^2 = E\int_0^{t_1 \wedge t_2} \left(Z_i - K(\beta_0', t_1 \wedge t_2, u)\right)^2 Y_i(t_1 \wedge t_2, u) e^{\beta_0 Z_i} \Lambda_0(du).$$

The limiting process in Theorem 17 has a complicated distribution that depends on unknown quantities including the baseline hazard rate and, in general, the value of β. In order to use this result one needs to be able to approximate the asymptotic significance points of the distribution.

The weighted bootstrap technique is a useful tool. We refer to [7] and [15] in the present setup when t is fixed.

Let $\varpi_1, \varpi_2, \ldots, \varpi_n$ be i.i.d. random variables, independent of $\tau_i, T_i, C_i, Z_i, i = 1, 2, \ldots n$, with zero mean and variance one. Define the processes $U_1^{\varpi}(t,u) = \sum_{i=1}^n N_i(t,u)\varpi_i$, $U_2^{\varpi}(t,s) = \sum_{i=1}^n \int_0^s Y_i(t,u)Z_i(u)N_i(t,\mathrm{d}u)\varpi_i$,

$$U^{\varpi}(\beta;t,s) = U_2^{\varpi}(t,s) - \int_0^s \overline{Z}(\beta;t,u)U_1^{\varpi}(t,\mathrm{d}u),$$

$$\begin{aligned}\widehat{\psi}_n^{\varpi}(z,t,s) &= \psi_n^{f\varpi}(z,t,s) - \int_0^t g_n(\widehat{\beta}(t,t),z,t,u)n^{-1/2}U_1^{\varpi}(t,\mathrm{d}u)\\ &\quad - \widehat{W}_n(\widehat{\beta}(t,t),z,t,s)U^{\varpi}(\widehat{\beta}(t,t),t,t),\end{aligned} \tag{20}$$

where, for $\widehat{\Lambda}$ defined by (14),

$$g_n(\beta,z,t,s) = \frac{\sum_{i=1}^n f(Z_i)\, I_{[Z_i\le z]}Y_i(t,s)\mathrm{e}^{\beta' Z_i}}{\sum_{i=1}^n I_{[Z_i\le z]}Y_i(t,s)\mathrm{e}^{\beta' Z_i}},$$

$$\begin{aligned}\widehat{W}_n(\beta,z,t,s) &= n^{-1}\mathcal{J}^{-1}(\beta,t)\sum_{k=1}^n \int_0^s Y_k(t,u)\mathrm{e}^{\beta' Z_k} f(Z_k)I_{[Z_k\le z]}\\ &\quad\times \left\{Z_k - \overline{Z}(\beta_0,t,u)\right\}\widehat{\Lambda}(\widehat{\beta}(t,t),t,\mathrm{d}u),\end{aligned}$$

$$\psi_n^{f\varpi}(z,t,s) = n^{-1/2}\sum_{i=1}^n f(Z_i)\, I_{[Z_i\le z]}\, N_i(t,s)\varpi_i. \tag{21}$$

Not only is (21) a strong martingale, the following theorem establishes that it is also a strong martingale given the data (τ_i, T_i, C_i, Z_i), $i = 1, \ldots, n$. Let $\mathcal{F}^{i\varpi}_{z,t,s} = \sigma\{\tau_i, T_i, C_i, Z_i, I_{[Z_i\le z']}N_i(u,v)\varpi_i, N_i(u,v)\varpi_i, 0\le u\le s, 0\le v\le t, a\le z'\le z\}$, and $\mathcal{F}^{\varpi *}_{z,t,s}$ be defined like (18), with $\mathcal{F}^{i\varpi}_{z,t,s}$ replacing $\mathcal{F}^{i}_{z,t,s}$.

Theorem 18. *The process $\psi_n^{f\varpi}$ of (21) is a strong martingale with respect to $\mathcal{F}^{\varpi *}_{z,t,s}$. Along almost all sample sequences, (τ_i, T_i, C_i, Z_i), $i = 1, \ldots$, given $\{(\tau_i, T_i, C_i, Z_i), i = 1, \ldots, n\}$, in the space $D([a,b)^p \times \widetilde{S}_*)$, $\psi_n^{f\varpi} \to_D \psi_\infty^f$.*

Along almost all sample sequences, (τ_i, T_i, C_i, Z_i), $i = 1, \ldots$, given $\{(\tau_i, T_i, C_i, Z_i), i = 1, \ldots, n\}$, in the space $D([a,b)^p \times \widetilde{S}_*)$, $\widehat{\psi}_n^{\varpi} \to_D \varsigma$.

If we replace the indicator function in (16) by $1_{[\widehat{\beta}(t,t)' Z_i \le v]}$, we obtain a process that can be used to test for the exponential link function. Similar to Theorem 16, we would obtain a three-dimensional strong martingale and convergence results similar to those obtained in Theorem 16 to Theorem 18 hold.

6. Proofs of the main theorems

6.1. Proofs of the strong martingale theorems of section 3

Proof of theorem 2. As defined, M_i is adapted to $\mathcal{F}_{t,s}$. To prove that $E(M_i(D)|\mathcal{F}^*_{t_0,s_0}) = 0$, a.s. over any rectangle $D = (t_0, t_0'] \times (s_0, s_0']$, we will show that $\sum_{j=1}^7 E(I_{E_{ij}} M_i(D)|\mathcal{F}^*_{t_0,s_0}) = 0$, a.s., where E_{ij} is defined by (23) below.

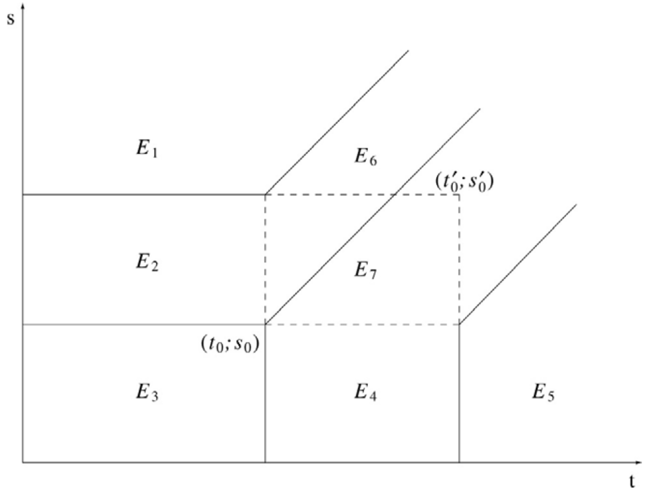

Fig. 1. Seven sets.

First we partition the first quadrant into seven disjoint sets E_1, E_2, E_3, E_4, E_5, E_6, and E_7, where

$$\begin{aligned}
E_1 &= \{(t,s) : 0 \le t \le s + (t_0 - s_0'), s > s_0'\},\\
E_2 &= \{(t,s) : 0 \le t \le t_0,\ s_0 < s \le s_0'\},\\
E_3 &= \{(t,s) : 0 \le t \le t_0,\ 0 \le s \le s_0\},\\
E_4 &= \{(t,s) : t_0 < t \le t_0',\ 0 \le s \le s_0\},\\
E_5 &= \{(t,s) : t > t_0',\ 0 \le s < t - (t_0' - s_0)\},\\
E_6 &= \{(t,s) : t > t_0,\ t - (t_0 - s_0) \le s < t - (t_0 - s_0')\},\\
E_7 &= \{(t,s) : s + (t_0 - s_0) < t \le s + (t_0' - s_0), s > s_0\}.
\end{aligned} \tag{22}$$

Please refer to Fig. 1 for a graphical portrayal of these seven sets. For individual i, the end point $(\widetilde{T}_i + \tau_i, \widetilde{T}_i)$ of its line in the Lexis diagram will fall into one of these seven sets. Therefore, for $j = 1, 2, \ldots, 7$, define the set E_{ij} as

$$E_{ij} = \{\omega : (\widetilde{T}_i(\omega) + \tau_i(\omega), \widetilde{T}_i(\omega)) \in E_j\}, \tag{23}$$

where E_j is defined by (22). Consequently, for each i, the sample space is partitioned into the seven sets, i.e. $\Omega = \cup_{j=1}^{7} E_{ij}$ and $E_{il} \cap E_{ik} = \phi$ for $l \neq k$.

For simplicity, throughout the following, we suppress the subscript 0 from the coordinates of the rectangle $(t_0, t_0'] \times (s_0, s_0']$. It is easy to establish that $I_{E_{ij}} M_i(D) = 0$, a.s. for $j = 1, 2, \ldots, 5$. For the remaining two cases, let $s^* = \sup\{u : Y_i(t,u) = 1\}$ and $s^{**} = \sup\{v : Y_i(t', v) = 1\}$. For each fixed t, s^* is an $\mathcal{F}_{t,s}$-stopping time, i.e. $[s^* < s] \in \mathcal{F}^i_{t,s} \subset \mathcal{F}_{t,s}$, for all $s \ge 0$ [6]. Since $\mathcal{F}^i_{t,s} \subseteq \mathcal{F}^i_{t',s}$, $[s^* < s] \in \mathcal{F}^i_{t',s} \subset \mathcal{F}_{t',s}$, for all $s \ge 0$. It follows that s^* is also an $\mathcal{F}_{t',s}$-stopping time. Let $\mathcal{F}_{t',s^*}$ consist of all sets $A \in \vee_{s\ge 0} \mathcal{F}_{t',s}$ such that $A \cap [s^* < s] \in \mathcal{F}_{t',s}$ for all $s \ge 0$, then $\mathcal{F}_{t',s^*}$ itself is a σ-field by T36, A1 in [6]. Heuristically, $\mathcal{F}_{t',s^*}$ contains all the information available by time s^*. Similarly, for fixed t', s^{**} is an $\mathcal{F}_{t',s}$-stopping time.

Then consider the case E_{i6}. Define two rectangles

$$D_1 = (t, t'] \times (s^*, s'] \quad \text{and} \quad D_2 = (t, t'] \times (s, s^*].$$

Note that $D_1 \cap D_2 = \phi$ and $D = D_1 \cup D_2$. Since $Y_i(v, u) = 1$ and $N_i(v, u) = 0$ for any $(v, u) \in D_2$, $M_{i,E_{i6}}(D_2) = 0$, a.s. Also, $I_{E_{i6}}[M_i(t, s^{**} \wedge s') - M_i(t, s^*)] = 0$, a.s. Thus, $I_{E_{i6}} M_i(D) = I_{E_{i6}} M_i((D_1)) = I_{E_{i6}} M_i(t', s^{**} \wedge s') - I_{E_{i6}} M_i(t', s^*)$, a.s. To prove $E[I_{E_{i6}} M_i(D)|\mathcal{F}^*_{t,s}] = 0$, a.s., we need to prove that for any set $G^* \in \mathcal{F}^*_{t,s}$,

$$\int_{G^*} I_{E_{i6}} M_i(D) \mathrm{d}P = 0. \tag{24}$$

The collection $\mathcal{G}$ of sets $\cap_{l=1}^n G_l$, where each G_l is of the form

$$\begin{aligned} G_l = \bigcap_{j=1}^{k} \{ & (I_{[\tau_l \le t_j]}, \tau_l I_{[\tau_l \le t_j]}, I_{[T_l \le s_j \wedge C_l \wedge (t_j - \tau_l)^+]}, T_l I_{[T_l \le s_j \wedge C_l \wedge (t_j - \tau_l)^+]}, \\ & I_{[C_l \le s_j \wedge T_l \wedge (t_j - \tau_l)^+]}, C_l I_{[C_l \le s_j \wedge T_l \wedge (t_j - \tau_l)^+]}, Z_l((s_j \wedge (t_j - \tau_l)^+)^+) I_{[\tau_l \le t_j]}) \in B_j \}, \end{aligned} \tag{25}$$

is a π-system that generates $\mathcal{F}^*_{t,s}$. Hence, we only need to prove that for any set $\cap_{l=1}^n G_l$, $\int_{\cap_{l=1}^n G_l} I_{E_{i6}} M_i(D) \mathrm{d}P = 0$. By Lemma 19, $G_i \cap E_{i6} \in \mathcal{F}^i_{t',s^*}$. By the independence of the sets $G_l \cap E_{i6}$ and $G_l, l \neq i$,

$$\begin{aligned} \int_{\cap_{l=1}^n G_l} I_{E_{i6}} M_i(D) \mathrm{d}P &= \int_{\cap_{l \neq i} G_l} \mathrm{d}P \int_{G_i \cap E_{i6}} M_i(D) \mathrm{d}P \\ &= \prod_{l \neq i} P(G_l) \int_{G_i \cap E_{i6}} [M_i(t', s^{**} \wedge s') - M_i(t', s^*)] \mathrm{d}P. \end{aligned}$$

Since for any given t', $M_i(t', s)$ is a one-dimensional martingale in s, and $s^{**} \wedge s'$ is a stopping time, therefore $E[M_i(t', s^{**} \wedge s') - M_i(t', s^*)|\mathcal{F}_{t',s^*}] = 0$, a.s., by the Optional Sampling Theorem [11]. That is, for $G_i \cap E_{i6} \in \mathcal{F}^i_{t',s^*}$, we have $\int_{G_i \cap E_{i6}} \left\{ M_i(t', s^{**} \wedge s') - M_i(t', s^*) \right\} \mathrm{d}P = 0$. So (24) is proven. Therefore, $E(I_{E_{i6}} M_i(D)|\mathcal{F}^*_{t,s}) = 0$, a.s.

Using Lemma 20 below, $E(I_{E_{i7}} M_i(D)|\mathcal{F}^*_{t,s}) = 0$, a.s., by an argument similar to that for $I_{E_{i6}} M_i(D)$. Hence Theorem 2 is proven. □

The following two lemmas were used in the above proof:

Lemma 19. *For $j = 1, 2, \ldots, k$, let $B_j = B_{j1} \times B_{j2} \times \cdots \times B_{j(p+6)}$ where $B_{jl}, l = 1, 2, \ldots, p+6$, is a rectangle in $[0, \infty)$ or $(-\infty, \infty)$. Let $\mathcal{G}_i$ be the collection of all sets of the form (25) with $l = i$, for some k, where $(t_j, s_j) \in [0, t] \times [0, \infty) \cup [0, \infty) \times [0, s]$. Then, for any $G_i \in \mathcal{G}_i$, there exists a G'_i such that*

$$E_{i6} \cap G_i = E_{i6} \cap G'_i, \tag{26}$$

where G'_i is a set of form

$$G_i' = \bigcap_{j=1}^{k}\{(I_{[\tau_i \le t_j']}, \tau_i I_{[\tau_i \le t_j']}, I_{[T_i \le s_j' \wedge C_i \wedge (t_j' - \tau_i)^+]}, T_i I_{[T_i \le s_j' \wedge C_i \wedge (t_j' - \tau_i)^+]},$$
$$I_{[C_i \le s_j' \wedge T_i \wedge (t_j' - \tau_i)^+]}, C_i I_{[C_i \le s_j' \wedge T_i \wedge (t_j' - \tau_i)^+]},$$
$$Z_i(s_j' \wedge (t_j' - \tau_i)^+) I_{[\tau_i \le t_j']}) \in B_j\},$$

and $(t_j', s_j') \preceq (t, s^*)$. *Hence* $E_{i6} \cap G_i \in \mathcal{F}^i_{t', s^*}$.

Lemma 19 states that for any generating set G_i of the strong past $\mathcal{F}^{i*}_{t,s}$, $E_{i6} \cap G_i = E_{i6} \cap G_i'$, where G_i' belongs to $\mathcal{F}^i_{t', s^*}$, the past of the stopping time s^*. Hence $E_{i6} \cap G_i$ belongs to $\mathcal{F}^i_{t', s^*}$.

Proof of lemma 19. Let X denote any of the seven generators of $\mathcal{F}^i_{t,s}$: $I_{[\tau_i \le t]}$, $\tau_i I_{[\tau_i \le t]}$, $I_{[T_i \le s \wedge C_i \wedge (t - \tau_i)^+]}$, $T_i I_{[T_i \le s \wedge C_i \wedge (t - \tau_i)^+]}$, $I_{[C_i \le s \wedge T_i \wedge (t - \tau_i)^+]}$, $C_i I_{[C_i \le s \wedge T_i \wedge (t - \tau_i)^+]}$, and $Z_i(s \wedge (t - \tau_i)^+) I_{[\tau_i \le t]}$. On the set E_{i6}, define G_i' as follows:

$$\begin{aligned} &t_j' = t_j, \quad s_j' = s_j, \quad \text{if } t_j \le t, s_j \le s^* \\ &t_j' = t_j, \quad s_j' = s^*, \quad \text{if } t_j \le t, s^* < s_j \\ &t_j' = t_j, \quad s_j' = s^*, \quad \text{if } t_j > t, s_j \le s. \end{aligned}$$

With G_i' so defined, the equation $X(t_j', s_j') = X(t_j, s_j)$ holds and hence (26) is proven. Since $E_{i6} \cap G_i \in \mathcal{F}^i_{t, s^*}$ and $\mathcal{F}^i_{t, s^*} \subset \mathcal{F}^i_{t', s^*}$, $E_{i6} \cap G_i \in \mathcal{F}^i_{t', s^*}$. □

Lemma 20. *Let* B_j, $j = 1, 2, \ldots, k$, *and* $\mathcal{G}_i$ *be defined as in* Lemma 19. *Then, for any* $G_i \in \mathcal{P}_i$, *there exists a* G_i' *such that*

$$E_{i7} \cap G_i = E_{i7} \cap G_i', \tag{27}$$

where G_i' *is a set of form*

$$G_i' = \bigcap_{j=1}^{k}\{(I_{[\tau_i \le t_j']}, \tau_i I_{[\tau_i \le t_j']}, I_{[T_i \le s_j' \wedge C_i \wedge (t_j' - \tau_i)^+]}, T_i I_{[T_i \le s_j' \wedge C_i \wedge (t_j' - \tau_i)^+]},$$
$$I_{[C_i \le s_j' \wedge T_i \wedge (t_j' - \tau_i)^+]}, C_i I_{[C_i \le s_j' \wedge T_i \wedge (t_j' - \tau_i)^+]},$$
$$Z_i(s_j' \wedge (t_j' - \tau_i)^+) I_{[\tau_i \le t_j']}) \in B_j\},$$

and $(t_j', s_j') \preceq (t', s)$. *Hence* $E_{i7} \cap G_i \in \mathcal{F}^i_{t', s}$.

Similar to Lemma 19, Lemma 20 states that for any generating set G_i of the strong past $\mathcal{F}^{i*}_{t,s}$, $E_{i7} \cap G_i \in \mathcal{F}^i_{t', s}$.

Proof of lemma 20. Let X denote any of N_i, Y_i, or $Y_i Z_i$. On the set E_{i7}, define G_i' as follows:

$$\begin{aligned} &t_j' = t_j, \quad s_j' = s_j, \quad \text{if } t_j \le t', \ s_j \le s \\ &t_j' = t_j, \quad s_j' = s, \quad \text{if } t_j \le t, \ s < s_j \\ &t_j' = t', \quad s_j' = s_j, \quad \text{if } t_j > t', \ s_j \le s. \end{aligned}$$

With G_i' so defined, the equation $X(t_j', s_j') = X(t_j, s_j)$ holds. Hence, (27) is proven. Consequently, $E_{i7} \cap G_i' \in \mathcal{F}^i_{t', s}$. □

Note that G_i defined as in (25) is closed under finite intersections. So $\mathcal{G}_i$ is a π-system generating the σ-field $\mathcal{F}^{i*}_{t,s}$ [5].

Proof of lemma 6. We have two cases.

(i) Suppose $\tau_i \le t-s$. If $\widetilde{T}_i \le s$, then $N_i(t',s) = N_i(t,s) = \delta_i$. For $u < \widetilde{T}_i$, $Y_i(t',u) = Y_i(t,u) = 1$. For $\widetilde{T}_i \le u$, $Y_i(t',u) = Y_i(t,u) = 0$. Hence $M_i(t',u) = M_i(t,u)$ for $u \le s$ and hence $J_{i2}(t',s) - J_{i2}(t,s) = 0$.

If $u < \widetilde{T}_i$, then $N_i(t',s) = N_i(t,s) = 0$ and $Y_i(t',u) = Y_i(t,u) = 1$, for $u \le s$. Again, $J_{i2}(t',s) - J_{i2}(t,s) = 0$.

(ii) Suppose $t-s < \tau_i$. Then, $Y_i(t,u) = 0$, for $s \le u \le s'$. If $(\tau_i + \widetilde{T}_i, \widetilde{T}_i)$ belongs to the triangle with vertices $(t-s,0)$, $(t,0)$, (t,s), then $N_i(t,s') = N_i(t,s) = \delta_i$; otherwise $N_i(t,s') = N_i(t,s) = 0$. In either case, $M_i(t',u) = M_i(t,u)$, for $s \le u \le s'$, and hence $J_{i1}(t,s') - J_{i1}(t,s) = 0$. Consequently, the Lemma 6 is proven. $\square$

Proof of lemma 7. For rectangle D of (4), recall (5). We have

$$
\begin{aligned}
&M_i^*(t',s') - M_i^*(t',s)\\
&\quad = \int_s^{s'} H_{i1}(u)Y_i(t',u)M_i(t',\mathrm{d}u)\int_0^s H_{i2}(v)Y_i(t',v)M_i(t',\mathrm{d}v)\\
&\qquad + \int_0^{s} H_{i1}(u)Y_i(t',u)M_i(t',\mathrm{d}u)\int_s^{s'} H_{i2}(v)Y_i(t',v)M_i(t',\mathrm{d}v)\\
&\qquad + \int_s^{s'} H_{i1}(u)Y_i(t',u)M_i(t',\mathrm{d}u)\int_s^{s'} H_{i2}(v)Y_i(t',v)M_i(t',\mathrm{d}v)\\
&\qquad - \left\{A_i^*(t',s') - A_i^*(t',s)\right\} = \sum_{j=1}^{4}\zeta_{1j}, \qquad \text{say.}
\end{aligned}
$$

Similarly, the increment $M_i^*(t,s') - M_i^*(t,s)$ equals

$$
\begin{aligned}
&\int_s^{s'} H_{i1}(u)Y_i(t,u)M_i(t,\mathrm{d}u)\int_0^s H_{i2}(v)Y_i(t,v)M_i(t,\mathrm{d}v)\\
&\qquad + \int_0^{s} H_{i1}(u)Y_i(t,u)M_i(t,\mathrm{d}u)\int_s^{s'} H_{i2}(v)Y_i(t,v)M_i(t,\mathrm{d}v)\\
&\qquad + \int_s^{s'} H_{i1}(u)Y_i(t,u)M_i(t,\mathrm{d}u)\int_s^{s'} H_{i2}(v)Y_i(t,v)M_i(t,\mathrm{d}v)\\
&\qquad - \left\{A_i^*(t,s') - A_i^*(t,s)\right\} = \sum_{j=1}^{4}\zeta_{2j}, \qquad \text{say.}
\end{aligned}
$$

Then, $\zeta_{11} - \zeta_{21}$ equals

$$
\begin{aligned}
&\int_s^{s'} H_1(u)Y_i(t',u)M_i(t',\mathrm{d}u)\int_0^s H_2(v)[Y_i(t',v)M_i(t',\mathrm{d}v) - Y_i(t,v)M_i(t,\mathrm{d}v)]\\
&\qquad + \int_s^{s'} H_1(u)[Y_i(t',u)M_i(t',\mathrm{d}u) - Y_i(t,u)M_i(t,\mathrm{d}u)]\int_0^s H_2(v)Y_i(t,v)M_i(t,\mathrm{d}v).
\end{aligned}
$$

By Lemma 6, the first term above equals zero. Since $\int_0^s H_2(v)Y_i(t,v)M_i(t,\mathrm{d}v)$ is measurable with respect to $\mathcal{F}^*_{t,s}$, by Theorem 3,

$$E(\zeta_{11}-\zeta_{21}|\mathcal{F}^*_{t,s}) = E\left(\int_s^{s'} H_1(u)\left[Y_i(t',u)M_i(t',\mathrm{d}u)-Y_i(t,u)M_i(t,\mathrm{d}u)\right]|\mathcal{F}^*_{t,s}\right) \times \int_0^s H_2(v)Y_i(t,v)M_i(t,\mathrm{d}v)=0 \text{ a.s.}$$

By the same argument, $E(\zeta_{12}-\zeta_{22}|\mathcal{F}^*_{t,s})=0$, a.s.

For the third terms, $\zeta_{13}-\zeta_{23}$ equals

$$\int_s^{s'} H_1(u)Y_i(t',u)M_i(t',\mathrm{d}u)\int_s^{s'} H_2(v)[Y_i(t',v)M_i(t',\mathrm{d}v)-Y_i(t,v)M_i(t,\mathrm{d}v)] + \int_s^{s'} H_1(u)[Y_i(t',u)M_i(t',\mathrm{d}u)-Y_i(t,u)M_i(t,\mathrm{d}u)]\int_s^{s'} H_2(v)Y_i(t,v)M_i(t,\mathrm{d}v).$$

We can write the first term above as

$$\begin{aligned}&\int_s^{s'} H_1(u)[Y_i(t',u)M_i(t',\mathrm{d}u)-Y_i(t,u)M_i(t,\mathrm{d}u)]\\ &\quad\times\int_s^{s'} H_2(v)[Y_i(t',v)M_i(t',\mathrm{d}v)-Y_i(t,v)M_i(t,\mathrm{d}u)]\\ &\quad+\int_s^{s'} H_1(u)Y_i(t,u)M_i(t,\mathrm{d}u)\int_s^{s'} H_2(v)[Y_i(t',v)M_i(t',\mathrm{d}v)-Y_i(t,v)M_i(t,\mathrm{d}v)].\end{aligned} \tag{28}$$

Since $\int_s^{s'} H_1(u)Y_i(t,u)M_i(t,\mathrm{d}u)$ is measurable with respect to $\mathcal{F}^*_{t,s}$, by Theorem 3, the conditional expectation of the second term of (28), given $\mathcal{F}^*_{t,s}$, equals zero, almost surely. Hence, we have $E(\zeta_{13}-\zeta_{23}|\mathcal{F}^*_{t,s}) = E(J_{1i}(D)J_{2i}(D)|\mathcal{F}^*_{t,s})$. Since $E(\zeta_{14}-\zeta_{24}|\mathcal{F}^*_{t,s}) = -E(A^*_i(D)|\mathcal{F}^*_{t,s})$, Lemma 7 is proven. □

Proof of theorem 4. Like the proof of Theorem 2, we consider the seven cases. By direct calculation, it is easy to establish that $E(I_{E_{ij}}M^*_i(D)|\mathcal{F}^*_{t,s}) = 0$ a.s. for $j=1,2,\ldots,5$. By Lemma 7,

$$\begin{aligned}I_{E_{i6}}M^*_i(D) &= I_{E_{i6}}\{J_{1i}(D)J_{2i}(D)-A^*_i(D)\}\\ &= I_{E_{i6}}\{[J_{1i}(t',s^{**}\wedge s')-J_{1i}(t',s^*)][J_{2i}(t',s^{**}\wedge s')-J_{2i}(t',s^*)]\\ &\quad -[A^*_i(t',s^{**}\wedge s')-A^*_i(t',s^*)]\}\\ &= I_{E_{i6}}[M^*_i(t',s^{**}\wedge s')-M^*_i(t',s^*)].\end{aligned}$$

Since $M^*_i(t',s)$ is an $\mathcal{F}^i_{t',s}$-martingale for fixed t', following the same argument as in the case of E_{i6} of Theorem 2, $E(M^*_{i,E_{i6}}(D)|\mathcal{F}^*_{t,s})=0$ a.s.

For case E_{i7}:

$$\begin{aligned}I_{E_{i7}}M^*_i(D) &= I_{E_{i7}}\{J_{1i}(D)J_{2i}(D)-A^*_i(D)\}\\ &= I_{E_{i7}}\{M^*_i(t',s^{**}\wedge s')-M^*_i(t',s).\}\end{aligned}$$

As in the proof of Theorem 2, $E(I_{E_{i7}}M^*_i(D)|\mathcal{F}^*_{t,s})=0$ a.s. □

Proof of lemma 9. From the definition of M, we know that M is a strong martingale. Since $A \geq 0$, increasing and continuous, we have $\sup_{(t,s)\leq(t^*,t^*)} |N(t,s)| \leq \sup_{(t,s)\leq(t^*,t^*)} |M(t,s)| + A(t^*,t^*)$. Hence

$$P\left[\sup_{(t,s)\leq(t^*,t^*)} |N(t,s)| \geq \varepsilon\right] \leq P\left[\sup_{(t,s)\leq(t^*,t^*)} |M(t,s)| \geq \frac{\varepsilon}{2}\right] + P\left[A(t^*,t^*) \geq \frac{\varepsilon}{2}\right].$$

Since $A(t,s) = \sum_{i=1}^n A_i(t,s)$ has continuous sample paths, the increment $A((t_-,s_-),(t,s)] = 0$, so that $P[\sup_{(t,s)\leq(t^*,t^*)} A((t_-,s_-),(t,s)] \leq \frac{\eta}{2}] = 1$, for any $\eta > 0$. Using Lemma 8, for the square-integrable strong martingale M,

$$\begin{aligned} P\left[\sup_{(t,s)\leq(t^*,t^*)} |N(t,s)| \geq \varepsilon\right] &\leq \frac{\eta}{\varepsilon^2} + P\left[A(t^*,t^*) \geq \frac{\eta}{4}\right] + P\left[A(t^*,t^*) \geq \frac{\varepsilon}{2}\right] \\ &\leq \frac{\eta}{\varepsilon^2} + 2P\left[A(t^*,t^*) \geq \frac{\eta}{4} \wedge \frac{\varepsilon}{2}\right]. \end{aligned}$$

The second statement follows from Lemma 8 since

$$\left(\sum_{i=1}^n \int_0^s Y_i(t,u)H_i(u)M_i(t,\mathrm{d}u)\right)^2 - \sum_{i=1}^n \int_0^s H_i^2(u)A_i(t,\mathrm{d}u)$$

is a strong martingale and $\sum_{i=1}^n \int_0^s H_i^2(u)A_i(t,\mathrm{d}u)$ is nondecreasing and continuous. □

6.2. *Proofs of Theorems 16–18*

Proof of theorem 16. Let $F_0^{(j)}$ denote the jth marginal of F_0 and let $\phi_0 : [a,b] \times (R^+)^3 \to [0,1]^p \times (R^+)^3$ be defined by $\phi_0(z_1,\ldots,z_p,x_1,x_2,x_3) = (F_0^{(1)}(z_1),\ldots,F_0^{(p)}(z_p),x_1,x_2,x_3)$. Then, ϕ_0 has an inverse $((F_0^{(1)})^{-1}(q_1),\ldots,(F_0^{(p)})^{-1}(q_p),x_1,x_2,x_3)$, where each $(F_0^{(j)})^{-1}(q_j)$ is the usual one-dimensional inverse. There exists a distribution function F_1 on $[0,1]^p \times (R^+)^3$, with its first p marginal distribution functions uniform on $[0,1]$, such that $F_0(z_1,\ldots,x_1,x_2,x_3) = F_1(\phi_0(z_1,\ldots,x_1,x_2,x_3))$. Let (U_i,C_i,T_i,τ_i) be an i.i.d. sequence with distribution function F_1, where U_i is p-dimensional, and let F_U denote the joint distribution function of U_i. Let $\phi : [a,b] \to [0,1]^p$ be defined by $\phi(z_1,\ldots,z_p) = (F_0^{(1)}(z_1),\ldots, F_0^{(p)}(z_p))$ with inverse ϕ^{-1}. Then, for $z \in [a,b]$, $F_U(\phi(z))$ is the distribution function of the Z_i. Hence

$$\begin{aligned} &\{\psi_n^f(z,t,s) : z \in [a,b],\ 0 \leq t,\ 0 \leq s,\ i = 1,\ldots,n\} \\ &\quad =_D \{\overline{\psi}_n^f(\phi(z),t,s) : z \in [a,b],\ 0 \leq t,\ 0 \leq s,\ i = 1,\ldots,n\}, \end{aligned}$$

where $Z_i = \phi^{-1}(U_i)$ and $\overline{\psi}_n^f(q,t,s) = n^{-1/2}\sum_{i=1}^n I_{[U_i\leq q]} f(\phi^{-1}(U_i))M_i(t,s)$. Then, each term of the above summation is a $(p+2)$-dimensional strong martingale and hence $\overline{\psi}_n^f$ is, also.

To prove weak convergence, the finite-dimensional distributions of $\overline{\psi}_n^f$ are sums of independent random vectors with zero mean and covariance given by

$$E\overline{\psi}_n^f(q,t,s)\overline{\psi}_n^f(q',t',s') = \overline{\Upsilon}_2(q \wedge q', t \wedge t', s \wedge s'), \tag{29}$$

where $\overline{\Upsilon}_2(q,t,s) = \int\int_{B(q,t,s,r)} [f(\phi^{-1}(q))]^2 h(u,q) \mathrm{e}^{\beta_0'\phi^{-1}(q)}\lambda_0(u)\mathrm{d}u F_U(\mathrm{d}q)F_\tau(\mathrm{d}r)$, F_τ is the distribution function of τ_i, $h(u,q) = E\left[I_{[u<T_i\wedge C_i]}|Z_i = \phi^{-1}(q)\right]$ and $B(q,t,s,r) = [0,q_1] \times \cdots \times [0,q_p] \times [0, s \wedge (t-r)^+]$. Hence the finite-dimensional distributions converge

to those of a Gaussian process $\overline{\psi}^f_\infty$ with covariance (29). Since $\overline{\Upsilon}_2$ is continuous, this process has a version with *a.s.* continuous sample paths and by Theorem 1, $\overline{\psi}^f_n \to_D \overline{\psi}^f_\infty$ and hence $\psi^f_n(z,t,s) \to_D \overline{\psi}^f_\infty(\phi(z),t,s)$. $\square$

Proof of theorem 17. Similar to the proof of Theorem 16, we can modify the process $\widehat{\psi}_n$ as follows: let

$$\begin{aligned}\overline{\widehat{\psi}}_n(q,t,s) &= \overline{\psi}^f_n(q,t,s) - \int_0^t \overline{g}(\beta,q,t,u) n^{-1/2} \sum_{i=1}^n M_i(t,\mathrm{d}u) \\ &\quad - \overline{W}_n(q,t,s)\widetilde{U}(\beta_0,t,t), \qquad (30)\end{aligned}$$

where $q \in [0,1]^p$ and $\overline{g}(\beta,q,t,u) = \overline{\Upsilon}_1(q,t,u)/\overline{\Upsilon}_0(b,t,u)$,

$$\begin{aligned}\overline{W}_n(q,t,s) &= n^{-1}\mathcal{J}^{-1}(\beta_0,t)\sum_{k=1}^n \int_0^s Y_k(t,u)\mathrm{e}^{\beta'\phi^{-1}(U_k)} f(\phi^{-1}(U_k)) I_{[U_k \le q]} \\ &\quad \times \left\{\phi^{-1}(U_k) - K(\beta_0,t,u)\right\} M_i(t,\mathrm{d}u).\end{aligned}$$

Convergence of the finite-dimensional distributions is straightforward. As to tightness, since the first two terms of (30) converge to Gaussian processes with continuous sample paths, by Lemma 6a of [23], we can prove tightness term by term. Lin et al. [15] proved tightness of the first and third terms of (30) in the case of non-staggered entries. However, their method is not appropriate for the second term. Our Theorem 16 implies tightness of the first term. For the second term, since $\psi^1_n \to_D \psi^1_\infty$, we can use the Skorohod–Dudley–Wichura representation theorem. On a single probability space, one can define $\widetilde{\psi}^1_n$ and $\widetilde{\psi}^1$ such that $\widetilde{\psi}^1_n =_D \overline{\psi}^1_n$, $\widetilde{\psi}^1 =_D \overline{\psi}^1_\infty$ and, almost surely, $\sup_{q,t,s} |\widetilde{\psi}^1_n(q,t,s) - \widetilde{\psi}^1(q,t,s)| \to 0$, a.s. Since $\overline{g}(\beta_0,q,t,u)$ is of bounded variation in u, uniformly in q and t, the difference $|\int_0^s \overline{g}(\beta_0,q,t,u)\widetilde{\psi}^1_n(b,t,\mathrm{d}u) - \int_0^s \overline{g}(\beta_0,q,t,u)\widetilde{\psi}^1(b,t,\mathrm{d}u)|$ converges to zero almost surely, uniformly in q, t and s. Thus, the second term of (30) converges in distribution and hence is tight.

Similar to the non-staggered case in [15], $W_n \to W$, almost surely, uniformly in q, t and s, while $\widetilde{U}_1(\beta_0,t,t)$ converges in distribution to $\xi(t,t)$ and hence the third term of (30) is tight. Consequently, $\overline{\widehat{\psi}}_n \to_D \overline{\varsigma}$ in the space $D([0,1]^p \times \widetilde{S}^*)$. Since $\overline{\widehat{\psi}}_n(\phi(z),t,s)$ and $\widehat{\psi}_n(z,t,s)$ have the same distribution, $\widehat{\psi}_n(z,t,s) \to_D \overline{\varsigma}(\phi(z),t,s)$, in the space $D([a,b)^p \times \widetilde{S}^*)$, where $\overline{\varsigma}(\phi(\cdot),\cdot,\cdot) =_D \varsigma(\cdot,\cdot,\cdot)$, and hence the theorem. $\square$

Proof of theorem 18. The process $\upsilon_i(z,t,s) = f(Z_i)\, I_{[Z_i \le z]} N_i(t,s)\varpi_i$ is clearly measurable with respect to $\mathcal{F}^{iw}_{z,t,s}$. Let D be the rectangle subtended by the points $x = (z,t,s)$ and $x' = (z',t',s')$ with $x \le x'$. Let $D_i = [t \le \tau_i + \widetilde{T}_i \le t', s \le \widetilde{T}_i \le s', T_i \le C_i, z \le Z_i \le z']$. Then, $I_{D_i^c}\upsilon_i(D) = 0$. Also, $E[I_{D_i}\upsilon_i(D)|\mathcal{F}^{\varpi *}_{z,t,s}] = I_{D_i} f(Z_i) E[\varpi_i|\mathcal{F}^{\varpi *}_{z,t,s}] = 0$ a.s., since the ϖ_i have zero mean and are independent of (τ_i, T_i, C_i, Z_i). Hence υ_i is a strong martingale and hence so is $\psi^{f\varpi}_n$.

Without loss of generality, we can assume that the Z_i have a continuous distribution function. (Otherwise, we can use the copula function ϕ of Theorems 16 and 17.) Given (τ_i, T_i, C_i, Z_i), $i = 1,2,\ldots,n$, the finite-dimensional distribution of $\psi^{f\varpi}_n$ is that of a sum of independent random variables with mean zero and covariance of the form $\sum_{i=1}^n [f(Z_i)]^2 I_{[Z_i \le z_1 \wedge z_2]} N_i(t_1 \wedge t_2, s_1 \wedge s_2)$, which converges *a.s.* to $\Upsilon_2(z_1 \wedge z_2, t_1 \wedge t_2, s_1 \wedge s_2)$. By the multivariate central limit theorem, its finite-dimensional distributions converge to those of ψ_∞, along almost all sample sequences,

(τ_i, T_i, C_i, Z_i), $i = 1, \ldots$, given $\{(\tau_i, T_i, C_i, Z_i), i = 1, \ldots, n\}$. Hence, by Theorem 1, the weak convergence of $\psi_n^{f\varpi}$ is proven. In a similar fashion, one can prove the last statement of Theorem 18. □

Acknowledgements

We wish to thank an anonymous referee for bringing the Ph.D. thesis of J. Hwang to our attention. The first author's research was partially supported by a Canadian NSERC discovery grant.

References

[1] O.O. Aalen, Statistical inference for a family of counting process, Ph.D. Dissertation, Univ. of California, Berkeley, 1975.
[2] P.K. Andersen, O. Borgan, R.D. Gill, N. Keiding, Statistical Models Based on Counting Processes, Springer-Verlag, New York, 1992.
[3] P.K. Andersen, R.D. Gill, Cox's regression model for counting processes: A large sample study, Ann. Statist. 10 (1982) 1100–1120.
[4] Y. Bilias, M. Gu, Z. Ying, Towards a general asymptotic theory for Cox model with staggered entry, Ann. Statist. 25 (1997) 662–682.
[5] P. Billingsley, Probability and Measure, 2nd ed., Wiley, New York, 1986.
[6] P. Brémaud, Point Processes and Queues: Martingale Dynamics, Springer-Verlag, New York, 1981.
[7] M.D. Burke, Testing regression models: A strong martingale approach, Fields Inst. Commun. 44 (2004) 351–365.
[8] R. Cairoli, J.B. Walsh, Stochastic integral in the plane, Acta Math. 1 34 (1975) 111–183.
[9] R.M. Dudley, Weak convergence of probabilities on non-separable metric spaces and empirical measures on Euclidean spaces, Illinois. J. Math. 10 (1966) 109–126.
[10] D. Feng, The Cox model with staggered entry, Ph.D. Thesis, Univ of Calgary, Calgary, Canada, 2002.
[11] T.R. Fleming, D.P. Harrington, Counting Processes and Survival Analysis, Wiley, New York, 1991.
[12] A.A. Gushchin, On the general theory of random fields in the plane, Russ. Math. Surveys 37 (6) (1982) 55–80.
[13] J. Hwang, Sequential estimation for the proportional hazards model in the presence of nuisance parameters, Ph.D. Thesis, Purdue University, West Lafayette, USA, 1992.
[14] B.G. Ivanoff, Stopping times and tightness for multiparameter martingales, Statist. Probab. Letters 28 (1996) 111–114.
[15] D.Y. Lin, L.J. Wei, Z. Ying, Checking the Cox model with cumulative sums of martingale-based residuals, Biometrika 80 (1993) 557–572.
[16] P. Meyer, Théorie élémentaire des processus à deux indices, in: H. Korezlioglu, G. Mazziotto, J. Szpirglas (Eds.), Processus Aléatoires à Deux Indices, in: Lect. Notes Math., vol. 863, Springer-Verlag, Berlin, 1981, pp. 1–39.
[17] Yu.S. Mishura, Sufficient conditions for relative compactness of measures corresponding to two-parameter strong martingale, Theory Probab. Math. Statist. 34 (1987) 117–125.
[18] D. Pollard, Empirical Processes: Theory and Applications, IMS, Hayward, CA, 1990.
[19] O. Pons, A test of independence for two censored survival times, Scand. J. Statist. 13 (1986) 173–185.
[20] T. Sellke, D. Siegmund, Sequential analysis of the proportional hazards model, Biometrika 70 (1983) 315–326.
[21] G.R. Shorack, J.A. Wellner, Empirical Processes with Applications to Statistics, Wiley, New York, 1986.
[22] D. Siegmund, Sequential Analysis: Tests and Confidence Intervals, Springer-Verlag, New York, 1985.
[23] M.L. Straf, Weak convergence of stochastic processes with several parameters, in: Proc. 6th Berkeley Symp. Math. Statist. Probab., vol. 2, University of California Press, 1972, pp. 182–221.